Mini-course on data-driven control, part 1

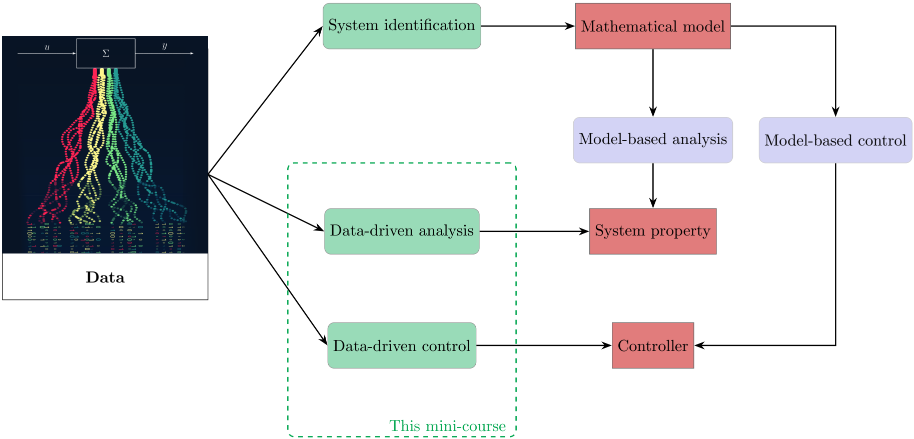
Henk van Waarde

Bernoulli Institute for Mathematics, Computer Science and AI
and
Jan C. Willems Center for Systems and Control

University of Groningen

Never a dull moment at the Benelux meeting...

Contents of the mini-course



Motivation:

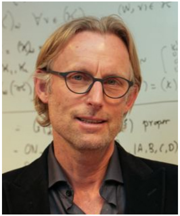
- We live in the era of big data
- Engineering systems are becoming more complex
- Direct approaches are promising in situations where modeling is challenging or computationally expensive

Contents of the mini-course, part 1

Acknowledgements

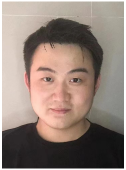
Kanat Camlibel

Jaap Eising



Harry Trentelman

Marieke Heidema (Tu-2.6-1)



Huayuan Huang (We-4.1-3)

Amir Shakouri (Tu-1.2-3)

Brayan Shali (We-6.3-4)

Jiwei Wang (Th-8.1-3)

Historical perspectives

A (coarse) timeline

1985
1986
1989
1999
2005
2008
2019
2020

Jan Willems.
From time series to linear system part II: exact modeling.

Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
On- and offline identification of linear state-space models.

Wouter Favoreel, Bart De Moor and Michel Gevers.
Subspace predictive control.

Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
A note on persistency of excitation.

Ivan Markovsky and Paolo Rapisarda.
Data-driven simulation and control.

Jeremy Coulson, John Lygeros and Florian Dörfler.
Data-enabled predictive control: in the shallows of the DeePC.

Claudio De Persis and Pietro Tesi.
Formulas for data-driven control: stabilization, optimality and robustness.

Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
Data informativity: a new perspective on data-driven analysis and control.

From 2020 onwards: a very quickly growing literature.

Preliminaries

Hankel matrix and persistency of excitation

Let $i, j \in \mathbb{N} = \{0, 1, \dots\}$ with $i \leqslant j$. Define $[i, j] := \{i, i+1, \dots, j\}$.
The restriction of $f : \mathbb{N} \to \mathbb{R}^q$ to $[i, j]$ is given by

$$f_{[i,j]} := \begin{bmatrix} f(i) \\ f(i+1) \\ \vdots \\ f(j) \end{bmatrix}.$$

Let k be a positive integer such that $k \leqslant j - i + 1$. Define the Hankel matrix of depth k, associated with $f_{[i,j]}$, as

$$\mathcal{H}_k(f_{[i,j]}) := \begin{bmatrix} f(i) & f(i+1) & \cdots & f(j-k+1) \\ f(i+1) & f(i+2) & \cdots & f(j-k+2) \\ \vdots & \vdots & & \vdots \\ f(i+k-1) & f(i+k) & \cdots & f(j) \end{bmatrix}.$$

Definition: $f_{[i,j]}$ is said to be persistently exciting of order k if $\mathcal{H}_k(f_{[i,j]})$ has full row rank (equal to kq).

Preliminaries

linear systems and behaviors

Consider the discrete-time linear time-invariant system

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned} \qquad \text{(LinSys)}$$

where $t \in \mathbb{N}$, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t) \in \mathbb{R}^p$.

The (input-output) behavior of (LinSys) is defined as the vector space

$$\mathfrak{B} := \left\{ (u,y) : \mathbb{N} \to \mathbb{R}^{m+p} \mid \exists\, x : \mathbb{N} \to \mathbb{R}^n \text{ s.t. (LinSys) holds } \forall t \in \mathbb{N} \right\}.$$

Given $t_1, t_2 \in \mathbb{N}$ with $t_2 \geqslant t_1$, we define the restricted behavior as

$$\mathfrak{B}_{[t_1,t_2]} := \left\{ \begin{bmatrix} u_{[t_1,t_2]} \\ y_{[t_1,t_2]} \end{bmatrix} \mid (u,y) \in \mathfrak{B} \right\}.$$

Facts:

1. By time-invariance, $\mathfrak{B}_{[t_1,t_2]} \supseteq \mathfrak{B}_{[t_1+1,t_2+1]} \supseteq \mathfrak{B}_{[t_1+2,t_2+2]} \supseteq \cdots$
2. If (A,B) is controllable then $\mathfrak{B}_{[t_1,t_2]} = \mathfrak{B}_{[t_1+1,t_2+1]} = \mathfrak{B}_{[t_1+2,t_2+2]} = \cdots$.

A (coarse) timeline

1985	
1986	Jan Willems. *From time series to linear system part II: exact modeling.*
1989	Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle. *On- and offline identification of linear state-space models.*
1999	Wouter Favoreel, Bart De Moor and Michel Gevers. *Subspace predictive control.*
2005	Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor. *A note on persistency of excitation.*
2008	Ivan Markovsky and Paolo Rapisarda. *Data-driven simulation and control.*
2019	Jeremy Coulson, John Lygeros and Florian Dörfler. *Data-enabled predictive control: in the shallows of the DeePC.*
2020	Claudio De Persis and Pietro Tesi. *Formulas for data-driven control: stabilization, optimality and robustness.*
2020	Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel. *Data informativity: a new perspective on data-driven analysis and control.*
2020	From 2020 onwards: a very quickly growing literature.

Subspace identification

problem formulation

Consider (LinSys) and assume that (A, B) is controllable and (C, A) is observable.

Problem: Given

- the data $\begin{bmatrix} u_{[0,T-1]} \\ y_{[0,T-1]} \end{bmatrix} \in \mathfrak{B}_{[0,T-1]}$, and
- an upper bound $N \geqslant n$

find matrices $\bar{A}, \bar{B}, \bar{C}$ and $\bar{D}$ such that

$$\bar{A} = SAS^{-1}, \;\; \bar{B} = SB, \;\; \bar{C} = CS^{-1}, \quad \text{and} \;\; \bar{D} = D$$

for some nonsingular matrix $S \in \mathbb{R}^{n \times n}$.

Thought experiment: Suppose that also $x_{[0,T]}$ is given.
Then (A, B, C, D) is a solution to the system of linear equations:

$$\begin{bmatrix} x(1) & x(2) & \cdots & x(T) \\ y(0) & y(1) & \cdots & y(T-1) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \\ u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}.$$

Subspace identification

state reconstruction from input-output data

The approach of Moonen et al.:
Assume $T \geqslant 2N$ and consider the partitioned Hankel matrix

$$\mathcal{H}_{2N}(u_{[0,T-1]}) = \left[\begin{array}{ccc} u(0) & \cdots & u(T-2N) \\ \vdots & & \vdots \\ u(N-1) & \cdots & u(T-N-1) \\ \hdashline u(N) & \cdots & u(T-N) \\ \vdots & & \vdots \\ u(2N-1) & \cdots & u(T-1) \end{array}\right] =: \left[\begin{array}{c} U_p \\ \hdashline U_f \end{array}\right].$$

Similarly, partition $\mathcal{H}_{2N}(y_{[0,T-1]})$ into the blocks Y_p and Y_f.

Define the state matrices (not given!):

$$\begin{aligned} X_p &:= \begin{bmatrix} x(0) & x(1) & \cdots & x(T-2N) \end{bmatrix} \\ X_f &:= \begin{bmatrix} x(N) & x(N+1) & \cdots & x(T-N) \end{bmatrix}. \end{aligned}$$

Subspace identification

state reconstruction from input-output data

Fact: If the matrix

$$\begin{bmatrix} X_p \\ \hdashline U_p \\ U_f \end{bmatrix} = \begin{bmatrix} \mathcal{H}_1(x_{[0,T-2N]}) \\ \hdashline \mathcal{H}_{2N}(u_{[0,T-1]}) \end{bmatrix}$$

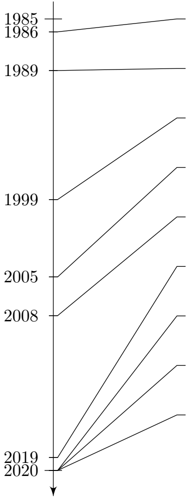
has full row rank then $\operatorname{rsp} X_f = \operatorname{rsp} \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \cap \operatorname{rsp} \begin{bmatrix} U_f \\ Y_f \end{bmatrix}$ and $\dim \operatorname{rsp} X_f = n$.

So a state sequence can be obtained by intersecting “past” and “future” data!

Remark: Idea of state reconstruction (using a “4 way infinite” Hankel matrix) used in 1986 by Willems (From times series paper part II).

An issue: The rank condition is not verifiable from data...

A (coarse) timeline



Jan Willems.
From time series to linear system part II: exact modeling.

Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
On- and offline identification of linear state-space models.

Wouter Favoreel, Bart De Moor and Michel Gevers.
Subspace predictive control.

Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
A note on persistency of excitation.

Ivan Markovsky and Paolo Rapisarda.
Data-driven simulation and control.

Jeremy Coulson, John Lygeros and Florian Dörfler.
Data-enabled predictive control: in the shallows of the DeePC.

Claudio De Persis and Pietro Tesi.
Formulas for data-driven control: stabilization, optimality and robustness.

Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
Data informativity: a new perspective on data-driven analysis and control.

From 2020 onwards: a very quickly growing literature.

Fundamental lemma

statement of the result

Theorem (Willems et al., 2005): Let $L \in [1, T]$. Assume that the pair (A, B) is controllable and that the input $u_{[0,T-1]}$ is persistently exciting of order $n + L$. Then the following statements hold:

1. $\operatorname{rank} \begin{bmatrix} \mathcal{H}_1(x_{[0,T-L]}) \\ \mathcal{H}_L(u_{[0,T-1]}) \end{bmatrix} = n + mL.$
2. $\operatorname{im} \begin{bmatrix} \mathcal{H}_L(u_{[0,T-1]}) \\ \mathcal{H}_L(y_{[0,T-1]}) \end{bmatrix} = \mathfrak{B}_{[0,L-1]}.$

Remarks:

- Rank conditions like item 1 are relevant for subspace identification ($L = 2N$).
- For persistency of excitation of order $n + L$ we need $T \geqslant (m+1)(n+L) - 1$ (sufficiently long trajectory).
- If $L > \ell$ where ℓ is the smallest integer such that

$$\operatorname{rank}\begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{\ell-1})^\top \end{bmatrix} = \operatorname{rank}\begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{\ell})^\top \end{bmatrix}$$

(i.e., the lag of the system), then $\mathfrak{B}_{[0,L-1]}$ uniquely determines $\mathfrak{B}$.

A (coarse) timeline

1985

- **1986** — Jan Willems. *From time series to linear system part II: exact modeling.*
- **1989** — Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle. *On- and offline identification of linear state-space models.*
- **1999** — Wouter Favoreel, Bart De Moor and Michel Gevers. *Subspace predictive control.*
- **2005** — Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor. *A note on persistency of excitation.*
- **2008** — Ivan Markovsky and Paolo Rapisarda. *Data-driven simulation and control.*
- **2019** — Jeremy Coulson, John Lygeros and Florian Dörfler. *Data-enabled predictive control: in the shallows of the DeePC.*
- **2020** — Claudio De Persis and Pietro Tesi. *Formulas for data-driven control: stabilization, optimality and robustness.*
- **2020** — Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel. *Data informativity: a new perspective on data-driven analysis and control.*
- **2020** — From 2020 onwards: a very quickly growing literature.

Data-driven tracking

problem statement

Problem: Let $L_{\text{ini}}, L_r \in \mathbb{N}$ be positive integers and define $L = L_{\text{ini}} + L_r$. Given:

- a symmetric positive semidefinite matrix

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \in \mathbb{R}^{(m+p)\times(m+p)},$$

- the data $\begin{bmatrix} u_{[0,T-1]} \\ y_{[0,T-1]} \end{bmatrix} \in \mathfrak{B}_{[0,T-1]}$,
- an initial trajectory $\begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L_{\text{ini}}-1]}$,
- and a reference signal $(v_{[L_{\text{ini}},L-1]}, z_{[L_{\text{ini}},L-1]}) \in \mathbb{R}^{mL_r} \times \mathbb{R}^{pL_r}$,

find a trajectory $(\bar{u}_{[L_{\text{ini}},L-1]}, \bar{y}_{[L_{\text{ini}},L-1]})$ that solves

$$\text{minimize} \sum_{t=L_{\text{ini}}}^{L-1} \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}^\top \Phi \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}$$

$$\text{subject to} \begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L-1]}.$$

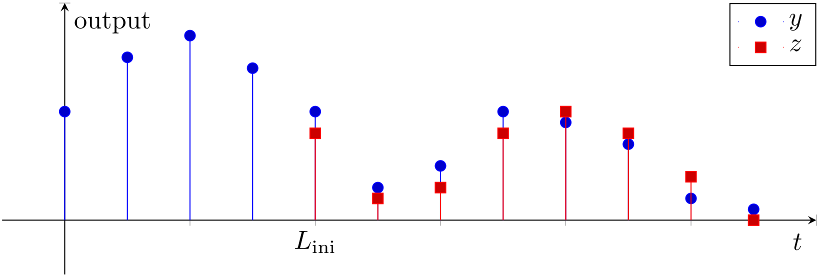
Data-driven tracking

problem statement

Problem: find a trajectory $(\bar{u}_{[L_{\text{ini}},L-1]}, \bar{y}_{[L_{\text{ini}},L-1]})$ that solves

$$\text{minimize} \sum_{t=L_{\text{ini}}}^{L-1} \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}^\top \Phi \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}$$

$$\text{subject to } \begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L-1]}.$$



Data-driven tracking

Notation: write $L = L_{\text{ini}} + L_r$ and partition

$$\mathcal{H}_L(u_{[0,T-1]}) = \begin{bmatrix} U_p \\ U_f \end{bmatrix} \text{ and } \mathcal{H}_L(y_{[0,T-1]}) = \begin{bmatrix} Y_p \\ Y_f \end{bmatrix}.$$

Under the assumptions of the fundamental lemma, the constraint is replaced by:

$$\begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \\ \bar{u}_{[L_{\text{ini}},L-1]} \\ \bar{y}_{[L_{\text{ini}},L-1]} \end{bmatrix} = \begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix} g \quad \text{for some} \quad g \in \mathbb{R}^{T-L+1}.$$

By defining $\tilde{\Phi} := \begin{bmatrix} I_{L_r} \otimes \Phi_{11} & I_{L_r} \otimes \Phi_{12} \\ I_{L_r} \otimes \Phi_{21} & I_{L_r} \otimes \Phi_{22} \end{bmatrix}$, the problem boils down to:

$$\begin{aligned} &\text{minimize} \quad \left\| \tilde{\Phi}^{\frac{1}{2}} \begin{bmatrix} U_f \\ Y_f \end{bmatrix} g - \tilde{\Phi}^{\frac{1}{2}} \begin{bmatrix} v_{[L_{\text{ini}},L-1]} \\ z_{[L_{\text{ini}},L-1]} \end{bmatrix} \right\|^2 \\ &\text{subject to} \quad \begin{bmatrix} U_p \\ Y_p \end{bmatrix} g = \begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \end{bmatrix}. \end{aligned} \qquad \text{(constrained LS)}$$

A (coarse) timeline

1985
1986
1989
1999
2005
2008
2019
2020

Jan Willems.
From time series to linear system part II: exact modeling.

Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
On- and offline identification of linear state-space models.

Wouter Favoreel, Bart De Moor and Michel Gevers.
Subspace predictive control.

Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
A note on persistency of excitation.

Ivan Markovsky and Paolo Rapisarda.
Data-driven simulation and control.

Jeremy Coulson, John Lygeros and Florian Dörfler.
Data-enabled predictive control: in the shallows of the DeePC.

Claudio De Persis and Pietro Tesi.
Formulas for data-driven control: stabilization, optimality and robustness.

Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
Data informativity: a new perspective on data-driven analysis and control.

From 2020 onwards: a very quickly growing literature.

Data-enabled predictive control

idea of receding horizon

Given: data, Φ, initial trajectory, and reference signal $\{(v(t), z(t))\}_{t=L_{\text{ini}}}^{\infty}$.

1. Set $\tau = 0$
2. Compute a trajectory $(\bar{u}_{[L_{\text{ini}}, L-1]}, \bar{y}_{[L_{\text{ini}}, L-1]})$ solving

$$\text{minimize} \sum_{t=L_{\text{ini}}}^{L-1} \begin{bmatrix} \bar{u}(t) - v(t+\tau) \\ \bar{y}(t) - z(t+\tau) \end{bmatrix}^{\top} \Phi \begin{bmatrix} \bar{u}(t) - v(t+\tau) \\ \bar{y}(t) - z(t+\tau) \end{bmatrix}$$

$$\text{subject to } \begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L-1]}.$$

3. Apply $\bar{u}(L_{\text{ini}})$ to the system, and measure $\bar{y}(L_{\text{ini}})$
4. Update $\bar{u}_{[0,L_{\text{ini}}-1]} \leftarrow \bar{u}_{[1,L_{\text{ini}}]}$ and $\bar{y}_{[0,L_{\text{ini}}-1]} \leftarrow \bar{y}_{[1,L_{\text{ini}}]}$
5. Set $\tau \leftarrow \tau + 1$ and go to step 2.

Remarks on generalizations:

- Constraints on inputs and outputs, regularization methods
- Stability analysis under terminal constraints[1].

[1] Berberich et al., *Data-driven model predictive control with stability and robustness guarantees*, IEEE TAC, 2020

Data-enabled predictive control

some remarks on model-based versions

Conceptually simple direct data-driven control scheme

Coulson et al. showed equivalence with model predictive control (MPC)

If the assumptions of the fundamental lemma hold and $L_{\text{ini}} \geqslant \ell$ then

$$\exists g : \begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \\ \bar{u}_{[L_{\text{ini}},L-1]} \\ \bar{y}_{[L_{\text{ini}},L-1]} \end{bmatrix} = \begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix} g \iff \bar{y}_{[L_{\text{ini}},L-1]} = Y_f \begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix}^{\dagger} \begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \\ \bar{u}_{[L_{\text{ini}},L-1]} \end{bmatrix}.$$

The subspace predictor $Y_f \begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix}^{\dagger}$ is a solution of $\min_K \left\| Y_f - K \begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix} \right\|_F$.

This predictor plays an important role in the MPC scheme of Favoreel et al.[2,3]

---

[2] Dörfler et al., "Bridging direct and indirect data-driven control formulations via regularizations and relaxations", *IEEE TAC*, 2023.

[3] Van Wingerden et al., "DeePC with instrumental variables: the direct equivalence with subspace predictive control", *IEEE CDC*, 2022.

A (coarse) timeline

- 1985
- 1986 — Jan Willems.
  From time series to linear system part II: exact modeling.
- 1989 — Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
  On- and offline identification of linear state-space models.
- 1999 — Wouter Favoreel, Bart De Moor and Michel Gevers.
  Subspace predictive control.
- 2005 — Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
  A note on persistency of excitation.
- 2008 — Ivan Markovsky and Paolo Rapisarda.
  Data-driven simulation and control.
- 2019 — Jeremy Coulson, John Lygeros and Florian Dörfler.
  Data-enabled predictive control: in the shallows of the DeePC.
- 2020 — Claudio De Persis and Pietro Tesi.
  Formulas for data-driven control: stabilization, optimality and robustness.
- 2020 — Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
  Data informativity: a new perspective on data-driven analysis and control.
- 2020 — From 2020 onwards: a very quickly growing literature.

Formulas for data-driven control

parameterization using G

Now consider the input-state dynamics $x(t+1) = Ax(t) + Bu(t)$, and the input-state data $X := \mathcal{H}_1(x_{[0,T]})$ and $U_- := \mathcal{H}_1(u_{[0,T-1]})$.

Define $X_+ := \mathcal{H}_1(x_{[1,T]})$ and $X_- := \mathcal{H}_1(x_{[0,T-1]})$. Note: $X_+ = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix}$.

Assume that (A, B) is controllable and $u_{[0,T-1]}$ is PE of order $n+1$. Then

$$\operatorname{rank} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = n + m.$$

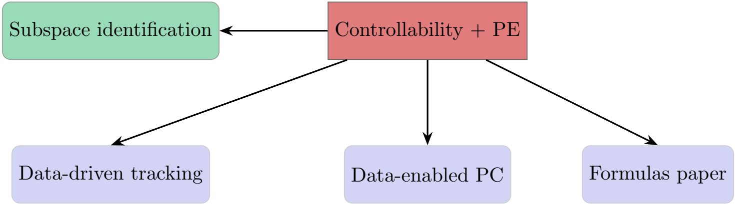
Thus, for any feedback controller $u = Kx$, there exists $G \in \mathbb{R}^{T \times n}$ such that

$$\begin{bmatrix} I \\ K \end{bmatrix} = \begin{bmatrix} X_- \\ U_- \end{bmatrix} G.$$

This leads to the data-based representation of the closed-loop system:

$$x(t+1) = (A + BK)x(t) = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} Gx(t) = X_+ Gx(t).$$

Questions



Important questions:

1. can we find controllers in situations where unique identification is impossible?
2. what is the minimum number of samples needed for control design?
3. and what about noisy data?

A (coarse) timeline

1985
1986
1989
1999
2005
2008
2019
2020

Jan Willems.
From time series to linear system part II: exact modeling.

Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
On- and offline identification of linear state-space models.

Wouter Favoreel, Bart De Moor and Michel Gevers.
Subspace predictive control.

Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
A note on persistency of excitation.

Ivan Markovsky and Paolo Rapisarda.
Data-driven simulation and control.

Jeremy Coulson, John Lygeros and Florian Dörfler.
Data-enabled predictive control: in the shallows of the DeePC.

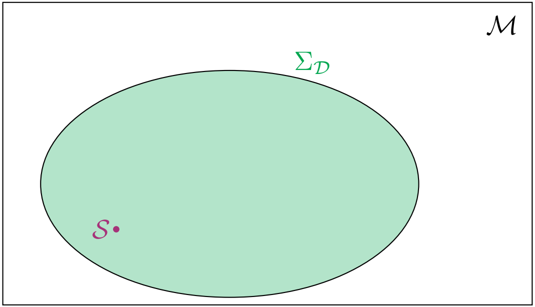
Claudio De Persis and Pietro Tesi.
Formulas for data-driven control: stabilization, optimality and robustness.

Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
Data informativity: a new perspective on data-driven analysis and control.

From 2020 onwards: a very quickly growing literature.

The point of view

informativity for data-driven analysis and control



$\mathcal{M}$: model class

$\mathcal{S}$: unknown system

$\mathcal{D}$: given data set

$\Sigma_{\mathcal{D}}$: data consistent systems

$\mathcal{P}$: system property

$\mathcal{O}$: control objective

- Data $\mathcal{D}$ are informative for
  - $\mathcal{P}$ $:\iff$ all systems in $\Sigma_{\mathcal{D}}$ have property $\mathcal{P}$
  - $\mathcal{O}$ $:\iff$ there exists a controller $\mathcal{C}$ that achieves $\mathcal{O}$ for all systems in $\Sigma_{\mathcal{D}}$.
- Data-driven control design := use $\mathcal{D}$ to find such $\mathcal{C}$
- Type of robust control problem, where
  - uncertainty stems from imperfect data (small number of samples, noisy, etc.)

Informativity for analysis and control

problems tackled so far

Problem	Data
controllability	E-IS
observability	E-S
stabilizability	E-IS
stability	E-S
state feedback stabilization	E-IS
deadbeat controller	E-IS
LQR	E-IS
suboptimal LQR	E-IS
suboptimal $\mathcal{H}_2$	E-IS
synchronization	E-IS
model reference control	E-IS
dynamic feedback stabilization	E-ISO
dynamic feedback stabilization	E-IO
dissipativity	E-ISO
tracking and regulation	E-IS
model reduction (moment matching)	E-IO
reachability (conic constraints)	E-IO

Problem	Data
stability	N-S
stabilizability	N-IS
state feedback stabilization	N-IS
state feedback $\mathcal{H}_2$ control	N-IS
state feedback $\mathcal{H}_\infty$ control	N-IS
model reference control	N-IS
dynamic feedback $\mathcal{H}_2$ control	N-IO
dynamic feedback $\mathcal{H}_\infty$ control	N-IO
stability	N-IO
dynamic feedback stabilization	N-IO
dissipativity	N-ISO
model reduction (balancing)	N-ISO
structural properties	N-ISO
absolute stabilization Lur'e systems	N-ISO

van Waarde *et al.*, "The Informativity Approach to Data-Driven Analysis and Control", *IEEE Control Systems Magazine*, 2023.

Controllability and stabilizability analysis

Controllability and stabilizability analysis

the setup

Consider the system $\mathcal{S}$ given by

$$x(t+1) = A_s x(t) + B_s u(t),$$

where $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$. $A_s \in \mathbb{R}^{n \times n}$ and $B_s \in \mathbb{R}^{m \times n}$ are unknown.

Data: $\mathcal{D} = (U_-, X)$ where $U_- := \mathcal{H}_1(u_{[0,T-1]})$ and $X := \mathcal{H}_1(x_{[0,T]})$.

Define $X_+ := \begin{bmatrix} x(1) & \cdots & x(T) \end{bmatrix}$ and $X_- := \begin{bmatrix} x(0) & \cdots & x(T-1) \end{bmatrix}$

Note that

$$X_+ = \begin{bmatrix} A_s & B_s \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix}.$$

No a priori assumptions on the data! All systems consistent with (U_-, X):

$$\Sigma_{\mathcal{D}} := \left\{ (A, B) \mid X_+ = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} \right\}.$$

Controllability and stabilizability analysis

the setup, definitions of informativity

Note that, in particular,

$$(A_s, B_s) \in \Sigma_{\mathcal{D}} := \left\{ (A,B) \mid X_+ = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} \right\}.$$

Definition: The data (U_-, X) are called informative for system identification if $\Sigma_{\mathcal{D}} = \{(A_s, B_s)\}$.

Fact: The data (U_-, X) are informative for system identification if and only if

$$\operatorname{rank} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = n + m.$$

Definition: We say the data (U_-, X) are informative for

- controllability if any $(A, B) \in \Sigma_{\mathcal{D}}$ is controllable.
- stabilizability if any $(A, B) \in \Sigma_{\mathcal{D}}$ is stabilizable.

Controllability and stabilizability analysis

data-driven Hautus tests

Theorem: The data (U_-, X) are informative for

- controllability if and only if $\operatorname{rank}(X_+ - \lambda X_-) = n \quad \forall \lambda \in \mathbb{C}$.
- stabilizability if and only if $\operatorname{rank}(X_+ - \lambda X_-) = n \quad \forall \lambda \in \mathbb{C}$ with $|\lambda| \geqslant 1$.

Example: Let $n = 2$, $m = 1$. Consider $U_- = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $X = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

$X_- = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $X_+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ so $X_+ - \lambda X_- = \begin{bmatrix} 1 & -\lambda \\ 0 & 1 \end{bmatrix}$.

The data (U_-, X) are thus informative for controllability.

However, (U_-, X) are not informative for system identification since

$$\Sigma_{\mathcal{D}} = \left\{ \left(\begin{bmatrix} 0 & a_1 \\ 1 & a_2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \mid a_1, a_2 \in \mathbb{R} \right\}.$$

Stabilization by state feedback

Stabilization by state feedback

definition, and "fundamental" lemma

Goal: Using the data (U_-, X), find a feedback law $u = Kx$ such that

$$x(t+1) = (A_s + B_s K)x(t)$$

is asymptotically stable (equivalently, $A_s + B_s K$ is Schur).

Definition: We say that the data (U_-, X) are informative for stabilization if there exists a feedback gain K such that $A + BK$ is Schur for all $(A, B) \in \Sigma_{\mathcal{D}}$.

Lemma: If the data (U_-, X) are informative for stabilization and K is a stabilizing feedback gain for all $(A, B) \in \Sigma_{\mathcal{D}}$ then

$$\operatorname{im} \begin{bmatrix} I \\ K \end{bmatrix} \subseteq \operatorname{im} \begin{bmatrix} X_- \\ U_- \end{bmatrix}.$$

Interpretation:

1. The matrix X_- necessarily has full row rank (excitation of all states).
2. Any stabilizing controller is of the form $K = U_- X_-^\sharp$, where $X_- X_-^\sharp = I$.

Stabilization by state feedback

main results

Theorem: The data (U_-, X) are informative for stabilization if and only if there exists a right inverse $X_-^\sharp$ of X_- such that $X_+X_-^\sharp$ is Schur. Moreover, K is stabilizing for all $(A,B) \in \Sigma_{\mathcal{D}}$ if and only if $K = U_-X_-^\sharp$, where $X_-^\sharp$ is as above.

Proof of "if" part: if $X_+X_-^\sharp$ is Schur and $K = U_-X_-^\sharp$ then for any $(A,B) \in \Sigma_{\mathcal{D}}$:

$$X_+X_-^\sharp = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} X_-^\sharp = A + BU_-X_-^\sharp.$$

Theorem: The data (U_-, X) are informative for stabilization if and only if there exists a $\Theta \in \mathbb{R}^{T \times n}$ such that

$$X_-\Theta = (X_-\Theta)^\top \quad \text{and} \quad \begin{bmatrix} X_-\Theta & X_+\Theta \\ \Theta^\top X_+^\top & X_-\Theta \end{bmatrix} > 0. \tag{LMI}$$

Moreover, K is a stabilizing feedback gain for all $(A,B) \in \Sigma_{\mathcal{D}}$ if and only if $K = U_-\Theta(X_-\Theta)^{-1}$ where Θ satisfies (LMI).

Remark: Same LMI condition as in De Persis & Tesi. But no PE requirements!

Stabilization by state feedback

main results

Theorem: The data (U_-, X) are informative for stabilization if and only if there exists a right inverse $X_-^\sharp$ of X_- such that $X_+X_-^\sharp$ is Schur. Moreover, K is stabilizing for all $(A,B) \in \Sigma_{\mathcal{D}}$ if and only if $K = U_-X_-^\sharp$, where $X_-^\sharp$ is as above.

Proof of "if" part: if $X_+X_-^\sharp$ is Schur and $K = U_-X_-^\sharp$ then for any $(A,B) \in \Sigma_{\mathcal{D}}$:

$$X_+X_-^\sharp = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} X_-^\sharp = A + BU_-X_-^\sharp.$$

Theorem: The data (U_-, X) are informative for stabilization if and only if there exist an $n \times n$ matrix $P = P^\top > 0$, an $L \in \mathbb{R}^{m \times n}$ and a scalar $\beta > 0$ such that

$$\begin{bmatrix} P - \beta I & 0 & 0 & 0 \\ 0 & -P & -L^\top & 0 \\ 0 & -L & 0 & L \\ 0 & 0 & L^\top & P \end{bmatrix} + \begin{bmatrix} X_+ \\ -X_- \\ -U_- \\ 0 \end{bmatrix} \begin{bmatrix} X_+ \\ -X_- \\ -U_- \\ 0 \end{bmatrix}^\top \geqslant 0. \qquad \text{(LMI2)}$$

Moreover, K is a stabilizing feedback gain for all $(A,B) \in \Sigma_{\mathcal{D}}$ if and only if $K = LP^{-1}$ where $P = P^\top > 0$ and L satisfy (LMI2) for some $\beta > 0$.

Stabilization by state feedback

comparison to identification

Example: Consider $n = 2$, $m = 1$ and the data

$$X = \begin{bmatrix} 1 & 0.5 & -0.25 \\ 0 & 1 & 1 \end{bmatrix}, \quad U_- = \begin{bmatrix} -1 & -1 \end{bmatrix}.$$

$X_+ X_-^{-1}$ is Schur and $K = U_- X_-^{-1} = \begin{bmatrix} -1 & -0.5 \end{bmatrix}$ is thus stabilizing. However,

$$\Sigma_{\mathcal{D}} = \left\{ \left(\begin{bmatrix} 1.5 + a_1 & 0.5a_1 \\ 1 + a_2 & 0.5 + 0.5a_2 \end{bmatrix}, \begin{bmatrix} 1 + a_1 \\ a_2 \end{bmatrix} \right) \mid a_1, a_2 \in \mathbb{R} \right\}.$$

The controller $K = U_- X_-^\sharp$ can be interpreted as a robust controller stabilizing all systems in $\Sigma_{\mathcal{D}}$.

For persistency of excitation of order $n + 1$ we need $T \geqslant n + m + nm$.
For identification of (A_s, B_s) we need $T \geqslant n + m$.
For stabilization we need $T \geqslant n$.

Conclusions Part 1

1. Historical perspectives
   - subspace identification
   - fundamental lemma
   - data-driven simulation and control
   - data-enabled predictive control
   - formulas for data-driven control
2. Data informativity
   - set of systems consistent with the data
   - analyze/control all systems in this set
   - informativity for stabilization weaker than for identification
   - not always the case (LQR)
3. Sets of systems become even more important in case of noise (Part 2!)

Thank you!

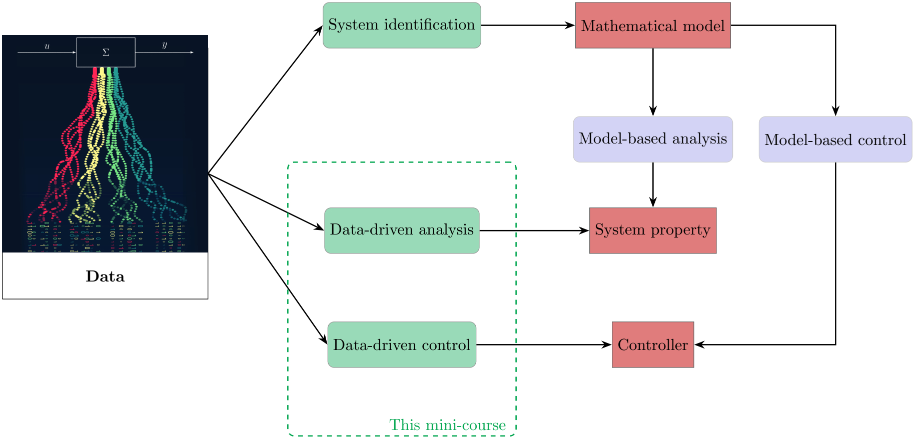
Mini-course on data-driven control, part 2

Henk van Waarde

Bernoulli Institute for Mathematics, Computer Science and AI
and
Jan C. Willems Center for Systems and Control

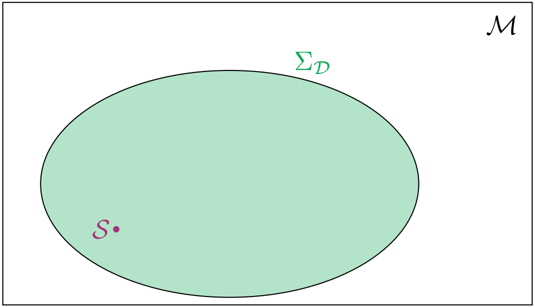
University of Groningen

Contents of the mini-course



The point of view

informativity for data-driven analysis and control



$\mathcal{M}$: model class

$\mathcal{S}$: unknown system

$\mathcal{D}$: given data set

$\Sigma_{\mathcal{D}}$: data consistent systems

$\mathcal{P}$: system property

$\mathcal{O}$: control objective

- Data $\mathcal{D}$ are informative for
  - $\mathcal{P} :\iff$ all systems in $\Sigma_{\mathcal{D}}$ have property $\mathcal{P}$
  - $\mathcal{O} :\iff$ there exists a controller $\mathcal{C}$ that achieves $\mathcal{O}$ for all systems in $\Sigma_{\mathcal{D}}$.
- Data-driven control design $:=$ use $\mathcal{D}$ to find such $\mathcal{C}$
- Type of robust control problem, where
  - uncertainty stems from imperfect data (small number of samples, noisy, etc.)

Informativity for analysis and control

problems tackled so far

Problem	Data
controllability	E-IS
observability	E-S
stabilizability	E-IS
stability	E-S
state feedback stabilization	E-IS
deadbeat controller	E-IS
LQR	E-IS
suboptimal LQR	E-IS
suboptimal $\mathcal{H}_2$	E-IS
synchronization	E-IS
model reference control	E-IS
dynamic feedback stabilization	E-ISO
dynamic feedback stabilization	E-IO
dissipativity	E-ISO
tracking and regulation	E-IS
model reduction (moment matching)	E-IO
reachability (conic constraints)	E-IO

Problem	Data
stability	N-S
stabilizability	N-IS
state feedback stabilization	N-IS
state feedback $\mathcal{H}_2$ control	N-IS
state feedback $\mathcal{H}_\infty$ control	N-IS
model reference control	N-IS
dynamic feedback $\mathcal{H}_2$ control	N-IO
dynamic feedback $\mathcal{H}_\infty$ control	N-IO
stability	N-IO
dynamic feedback stabilization	N-IO
dissipativity	N-ISO
model reduction (balancing)	N-ISO
structural properties	N-ISO
absolute stabilization Lur'e systems	N-ISO

van Waarde *et al.*, "The Informativity Approach to Data-Driven Analysis and Control", *IEEE Control Systems Magazine*, 2023.

Contents of the mini-course, part 2

1. Stabilization using noisy input-state data
2. Matrix versions of the S-lemma
3. Solution to stabilization problem
4. Performance: $\mathcal{H}_2$ control
5. Extension to input-output data
6. Conclusions

Stabilization using noisy data

Stabilization using noisy data

problem setup

Consider the system

$$x(t+1) = A_s x(t) + B_s u(t) + w(t),$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input and $w \in \mathbb{R}^n$ is the process noise.

The matrices A_s and B_s are unknown but the following data are given:

$$\begin{aligned} X &:= \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \\ U_- &:= \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}. \end{aligned}$$

Goal: using the data $\mathcal{D} = (U_-, X)$, find a feedback law $u = Kx$ such that

$$x(t+1) = (A_s + B_s K)x(t)$$

is asymptotically stable (equivalently, $A_s + B_s K$ is Schur).

Stabilization using noisy data

assumption on the noise

The matrix

$$W_- = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \end{bmatrix}$$

is unknown but assumed to satisfy the quadratic matrix inequality:

$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}}_{=:\Phi} \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0, \qquad \text{(QMI)}$$

for a known matrix $\Phi \in \mathbb{S}^{n+T}$.

Special cases:

- Energy bound: If $\Phi_{12} = 0$ and $\Phi_{22} = -I$ then $\sum_{t=0}^{T-1} w(t)w(t)^\top \leqslant \Phi_{11}$.
- Noiseless case: If, in addition, $\Phi_{11} = 0$ then $W_- = 0$.
- Sample bounds: If $||w(t)||^2 \leqslant \epsilon \; \forall t$ then energy bound holds with $\Phi_{11} = \epsilon T I$.
- Sample covariance bounds: If $\Phi_{22} = \frac{1}{T}(-I + \frac{1}{T}\mathbb{1}\mathbb{1}^\top)$, $\Phi_{12} = 0$ and $\mu = \frac{1}{T}\sum_{t=0}^{T-1} w(t)$ then

$$\frac{1}{T}\sum_{t=0}^{T-1}(w(t)-\mu)(w(t)-\mu)^\top \leqslant \Phi_{11}.$$

Stabilization using noisy data

some preliminaries on quadratic matrix inequalities

We will consider the set $\mathbf{\Pi}_{q,r}$, defined as

$$\left\{ \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} \in \mathbb{S}^{q+r} \mid \Pi_{22} \leqslant 0,\ \ker \Pi_{22} \subseteq \ker \Pi_{12},\ \Pi_{11} - \Pi_{12}\Pi_{22}^{\dagger}\Pi_{21} \geqslant 0 \right\}.$$

For $\Pi \in \mathbf{\Pi}_{q,r}$, define the set

$$\mathcal{Z}_r(\Pi) := \left\{ Z \in \mathbb{R}^{r \times q} \mid \begin{bmatrix} I \\ Z \end{bmatrix}^\top \Pi \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0 \right\}.$$

We assume that $\Phi \in \mathbf{\Pi}_{n,T}$. The assumption on the noise thus becomes:

$$W_-^\top \in \mathcal{Z}_T(\Phi).$$

Interpretation: $\Phi \in \mathbf{\Pi}_{n,T}$ implies that $\mathcal{Z}_T(\Phi)$ is convex and nonempty.

Stabilization using noisy data

informativity for stabilization

Introduce the matrices:

$$X_- := \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix}, \quad X_+ := \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}.$$

The set $\Sigma_{\mathcal{D}}$ of systems consistent with the data is given by:

$$\Sigma_{\mathcal{D}} = \{(A, B) \mid X_+ = AX_- + BU_- + W_- \text{ for some } W_-^\top \in \mathcal{Z}_T(\Phi)\}.$$

Definition: The data (U_-, X) are called informative for quadratic stabilization if there exists a feedback gain K and a matrix $P = P^\top > 0$ such that

$$P - (A + BK)P(A + BK)^\top > 0$$

for all $(A, B) \in \Sigma_{\mathcal{D}}$.

Problem: Find conditions for informativity, and provide a K (if it exists).

Stabilization using noisy data

implications of QMI's

All data-consistent systems satisfy a quadratic matrix inequality (QMI):

$$\Sigma_{\mathcal{D}} = \left\{ (A,B) \mid \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \right\}.$$

The Lyapunov inequality for stability of $A+BK$ can be written as:

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0,$$

which is also a quadratic matrix inequality in (A,B).

An important question:

When does one QMI hold for all (A,B) satisfying another QMI?

Classical S-lemma

Classical S-lemma

strict S-lemma and remarks

Theorem[1]: Let $M, N \in \mathbb{S}^n$. Then $x^\top Mx > 0$ for all nonzero $x \in \mathbb{R}^n$ satisfying $x^\top Nx \geqslant 0$ if and only if there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N > 0$.

V.A. Yakubovich
Russian control theorist
(1926 – 2012)

Remarks:

- Originally developed in the context of absolute stability of Lur'e systems.
- Non-strict version (assuming $N \nleq 0$).
- Finsler's lemma[2]: $x \neq 0$ and $x^\top Nx = 0 \implies x^\top Mx > 0$ if and only if $\exists\, \alpha \in \mathbb{R}$ such that $M - \alpha N > 0$.

But for our purposes we need a matrix version of the S-lemma!

[1] V. A. Yakubovich, "S-procedure in nonlinear control theory," *Vestnik Leningrad University Mathematics*, 1977.

[2] P. Finsler, Über das vorkommen definiter und semidefiniter formen in scharen quadratischer formen, *Comment. Math. Helv.*, 1936.

The matrix S-lemma

The matrix S-lemma

with α and β, and version for $N_{22} < 0$

Let $M, N \in \mathbb{S}^{q+r}$.

Define

$$\mathcal{Z}_r^+(M) := \left\{ Z \in \mathbb{R}^{r \times q} \mid \begin{bmatrix} I \\ Z \end{bmatrix}^\top M \begin{bmatrix} I \\ Z \end{bmatrix} > 0 \right\}.$$

Theorem[3]: Assume that $N \in \mathbf{\Pi}_{q,r}$ and $M_{22} \leqslant 0$. Then $\mathcal{Z}_r(N) \subseteq \mathcal{Z}_r^+(M)$ if and only if there exist scalars $\alpha \geqslant 0$ and $\beta > 0$ such that

$$M - \alpha N \geqslant \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix}.$$

Theorem[3]: Assume that $N \in \mathbf{\Pi}_{q,r}$ and $N_{22} < 0$. Then $\mathcal{Z}_r(N) \subseteq \mathcal{Z}_r^+(M)$ if and only if there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N > 0$.

Remark: Also matrix versions of non-strict S-lemma and Finsler's lemma!

[3] van Waarde *et al.*, "Quadratic Matrix Inequalities with Applications to Data-Based Control", *SICON*, 2023.

Application to data-driven stabilization

Application to data-driven stabilization

recall definition of informativity

The data (U_-, X) are informative for quadratic stabilization if there exist $P = P^\top > 0$ and K such that all systems in

$$\Sigma_{\mathcal{D}} = \left\{ (A,B) \mid \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} I & X_+ \\ \hdashline 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ \hdashline 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top}_{=:\, N} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \right\}$$

satisfy the Lyapunov inequality

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \underbrace{\left[\begin{array}{c:cc} P & 0 & 0 \\ \hdashline 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{array}\right]}_{=:\, M(P,K)} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0.$$

So the question is: when do there exist $P = P^\top > 0$ and K such that

$$\mathcal{Z}_r(N) \subseteq \mathcal{Z}_r^+(M(P,K))?$$

Application to data-driven stabilization

main theorem

Note that $N_{22} = \begin{bmatrix} X_- \\ U_- \end{bmatrix} \Phi_{22} \begin{bmatrix} X_- \\ U_- \end{bmatrix}^\top \leqslant 0$ and $M_{22}(P,K) = -\begin{bmatrix} I \\ K \end{bmatrix} P \begin{bmatrix} I \\ K \end{bmatrix}^\top \leqslant 0.$

Also, $\ker N_{22} \subseteq \ker N_{12}$, and $N_{11} - N_{12} N_{22}^\dagger N_{21} \geqslant 0$, thus we have $N \in \mathbf{\Pi}_{n,n+m}$.

Theorem: The data (U_-, X) are informative for quadratic stabilization if and only if there exist matrices $P = P^\top > 0$ and K and scalars $\alpha \geqslant 0$ and $\beta > 0$ such that

$$\begin{bmatrix} P - \beta I & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^\top & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \geqslant 0.$$

The matrix K stabilizes all systems in $\Sigma_{\mathcal{D}}$.

Application to data-driven stabilization

main theorem

Note that $N_{22} = \begin{bmatrix} X_- \\ U_- \end{bmatrix} \Phi_{22} \begin{bmatrix} X_- \\ U_- \end{bmatrix}^\top \leqslant 0$ and $M_{22}(P,K) = -\begin{bmatrix} I \\ K \end{bmatrix} P \begin{bmatrix} I \\ K \end{bmatrix}^\top \leqslant 0.$

Also, $\ker N_{22} \subseteq \ker N_{12}$, and $N_{11} - N_{12}N_{22}^\dagger N_{12} \geqslant 0$, thus we have $N \in \mathbf{\Pi}_{n,n+m}$.

Since α is necessarily positive, P and β can be scaled by $\frac{1}{\alpha}$. Then a Schur complement argument and change of variables lead to:

Theorem: The data (U_-, X) are informative for quadratic stabilization if and only if there exist matrices $P = P^\top > 0$ and L, and a scalar $\beta > 0$ such that

$$\begin{bmatrix} P-\beta I & 0 & 0 & 0 \\ 0 & -P & -L^\top & 0 \\ 0 & -L & 0 & L \\ 0 & 0 & L^\top & P \end{bmatrix} - \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix}^\top \geqslant 0.$$

The feedback gain $K = LP^{-1}$ stabilizes all systems in $\Sigma_{\mathcal{D}}$.

Application to data-driven stabilization

a conceptual algorithm

- Setting: model class of systems (A, B) with noise model given by Φ.
- Data: input/state samples (U_-, X).
- Check feasibility of the LMI: solutions $P = P^\top > 0$ of size $n \times n$, L of size $m \times n$, scalar $\beta > 0$.
- The feedback gain $K := LP^{-1}$ quadratically stabilizes all $(A, B) \in \Sigma_{\mathcal{D}}$.

Compared to related results in [3,4]:

- non-conservative control design
  - if and only if conditions for informativity via matrix S-lemma
- low complexity
  - Dimensions of the unknown matrices are independent of T
  - Size of the LMI is $(3n + m) \times (3n + m)$, also independent of T.

Compared to the results in [5]:

- no (generalized) Slater condition required!

[3] J. Berberich, A. Koch, C. W. Scherer, F Allgöwer, "Robust data-driven state-feedback design", *ACC*, 2020.

[4] C. De Persis and P. Tesi, "Formulas for Data-driven Control: Stabilization, Optimality and Robustness", *IEEE TAC*, 2020.

[5] van Waarde, Camlibel, and Mesbahi, "From Noisy Data to Feedback Controllers: Nonconservative Design via a Matrix S-Lemma", *IEEE TAC*, 2022.

Application to data-driven stabilization

the special case of noiseless data

Recall that the noiseless case ($W_- = 0$) can be captured by setting

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix}.$$

So the LMI now becomes:

$$\begin{bmatrix} P - \beta I & 0 & 0 & 0 \\ 0 & -P & -L^\top & 0 \\ 0 & -L & 0 & L \\ 0 & 0 & L^\top & P \end{bmatrix} + \begin{bmatrix} X_+ \\ -X_- \\ -U_- \\ 0 \end{bmatrix} \begin{bmatrix} X_+ \\ -X_- \\ -U_- \\ 0 \end{bmatrix}^\top \geqslant 0.$$

- Same LMI condition as for informativity for stabilization (Part 1!)
- So for noise-free data, informativity for stabilization and quadratic stabilization are equivalent.
- In this case, a common Lyapunov function is not conservative.

Data-driven $\mathcal{H}_2$ control

$\mathcal{H}_2$ control using noisy data

Some background on $\mathcal{H}_2$ suboptimal control

Consider the system

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + Du(t). \end{aligned}$$

The state feedback control law

$$u(t) = Kx(t)$$

yields the closed loop system

$$\begin{aligned} x(t+1) &= (A+BK)x(t) + w(t), \\ y(t) &= (C+DK)x(t). \end{aligned}$$

Closed loop impulse response:

$$G_K(t) := (C+DK)(A+BK)^t.$$

The $\mathcal{H}_2$ performance is given by: $J_{\mathcal{H}_2}(K) := \sum_{t=0}^{\infty} \operatorname{tr}(G_K^\top(t) G_K(t))$.

$\mathcal{H}_2$ control using noisy data

Some background on $\mathcal{H}_2$ suboptimal control

$\mathcal{H}_2$ suboptimal control problem: given a tolerance $\gamma > 0$, find a gain K such that $A + BK$ is Schur and $J_{\mathcal{H}_2}(K) < \gamma$.

Proposition: Let $\gamma > 0$ and K be given. Then $A + BK$ is Schur and $J_{\mathcal{H}_2}(K) < \gamma$ if and only if there exists $P = P^\top > 0$ such that

$$P > (A + BK)^\top P(A + BK) + (C + DK)^\top (C + DK) \tag{2}$$

and

$$\mathrm{tr}(P) < \gamma.$$

Put $Y := P^{-1}$ and $L = KY$. Then (2) is equivalent to:

$$Y - (CY + DL)^\top (CY + DL) > 0$$

and

$$\begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix}^\top \underbrace{\left[\begin{array}{c:c} Y & 0 \\ \hdashline 0 & -\begin{bmatrix} Y \\ L \end{bmatrix} (Y - C_{Y,L}^\top C_{Y,L})^{-1} \begin{bmatrix} Y \\ L \end{bmatrix}^\top \end{array}\right]}_{=:M(Y,L)} \begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix} > 0.$$

$\mathcal{H}_2$ control using noisy data

system and noise

Again consider an unknown system

$$x(t+1) = A_s x(t) + B_s u(t) + w(t).$$

In addition, a performance output

$$y(t) = Cx(t) + Du(t)$$

is specified by the designer, C and D are known matrices.

Again, for fixed T, unknown noise samples $w(0), w(1), \dots, w(T-1)$ affect the system. We assume that the possible noise sample matrices

$$W_- := \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \end{bmatrix}$$

satisfy the bound

$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \Phi \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0$$

for a known $\Phi \in \mathbf{\Pi}_{n,T}$.

$\mathcal{H}_2$ control using noisy data

definition of informativity

Consider the data: $\mathcal{D} = (U_-, X)$.

The set $\Sigma_{\mathcal{D}}$ of all systems consistent with the data:

$$\Sigma_{\mathcal{D}} := \left\{ (A,B) \mid \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top}_{=:N} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \right\}.$$

Goal of data-driven suboptimal $\mathcal{H}_2$ control: given $\gamma > 0$, find a gain K such that $A + BK$ is Schur and $J_{\mathcal{H}_2}(K) < \gamma$ for all $(A,B) \in \Sigma_{\mathcal{D}}$.

Definition: The data (U_-, X) are informative for $\mathcal{H}_2$ control with tolerance γ if there exists $P = P^\top > 0$ and a gain K such that $\operatorname{tr}(P) < \gamma$ and

$$P > (A+BK)^\top P(A+BK) + (C+DK)^\top (C+DK)$$

for all $(A,B) \in \Sigma_{\mathcal{D}}$.

$\mathcal{H}_2$ control using noisy data

implication of QMIs

Recall: putting $Y := P^{-1}$ and $L = KY$ and $C_{Y,L} := CY + DL$, this is equivalent to: $\exists\ Y > 0$ and L such that

$$Y - C_{Y,L}^\top C_{Y,L} > 0, \ \ \mathrm{tr}(Y^{-1}) < \gamma,$$

and

$$\begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix}^\top \underbrace{\left[\begin{array}{c:c} Y & 0 \\ \hdashline 0 & -\begin{bmatrix} Y \\ L \end{bmatrix} (Y - C_{Y,L}^\top C_{Y,L})^{-1} \begin{bmatrix} Y \\ L \end{bmatrix}^\top \end{array}\right]}_{=: M(Y,L)} \begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix} > 0$$

for all (A, B) that satisfy the QMI

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top}_{=: N} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0.$$

$\mathcal{H}_2$ control using noisy data

the main result

Question: When do there exist $Y > 0$ and L such that all solutions (A, B) of the first QMI (with N) also satisfy the second QMI (with $M(Y, L)$)?

We again solve this using the matrix S-lemma:

Theorem: The data (U_-, X) are informative for $\mathcal{H}_2$ control with performance γ if and only if there exist $Y = Y^\top > 0$, $L \in \mathbb{R}^{m \times n}$, $\alpha \geqslant 0$ and $\beta > 0$ such that $\operatorname{tr}(Y^{-1}) < \gamma$, $Y - C_{Y,L}^\top C_{Y,L} > 0$ and

$$\begin{bmatrix} Y - \beta I & 0 \\ 0 & -\begin{bmatrix} Y \\ L \end{bmatrix} (Y - C_{Y,L}^\top C_{Y,L})^{-1} \begin{bmatrix} Y \\ L \end{bmatrix}^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \geqslant 0.$$

$K := LY^{-1}$ is such that $A + BK$ is Schur and $J_{\mathcal{H}_2}(K) < \gamma$ for all $(A, B) \in \Sigma_{\mathcal{D}}$.

Remark: The inequalities can again be reformulated as LMIs.

Stabilization using input-output data

Stabilization using input-output data

the model class and data

Model class of all of autoregressive (AR) systems of known order L, input dimension m, output dimension p:

$$P(\sigma)y = Q(\sigma)u + w,$$

where $w(t)$ is unknown additive noise, σ is the shift operator $((\sigma f)(t) = f(t+1))$ and P, Q are polynomial matrices:

$$\begin{aligned} P(\xi) &= I\xi^L + P_{L-1}\xi^{L-1} + \cdots + P_1\xi + P_0, \\ Q(\xi) &= Q_{L-1}\xi^{L-1} + \cdots + Q_1\xi + Q_0, \end{aligned}$$

with unknown coefficients $P_0, P_1, \ldots, P_{L-1} \in \mathbb{R}^{p\times p}$, $Q_0, Q_1, \ldots, Q_{L-1} \in \mathbb{R}^{p\times m}$.

We collect input-output data

$$u(0), u(1), \ldots, u(T),\ \ y(0), y(1), \ldots, y(T),$$

for $T \geqslant L$, generated by the (unknown) true system

$$P_s(\sigma)y = Q_s(\sigma)u + w$$

within the model class.

Stabilization using input-output data

assumption on the noise

The $\tau := T - L + 1$ noise samples $w(0), w(1), \dots, w(T-L)$ are unknown.
However, we assume that the real $p \times \tau$ matrix

$$W := \begin{bmatrix} w(0) & w(1) & \cdots & w(T-L) \end{bmatrix},$$

satisfies the quadratic matrix inequality

$$\begin{bmatrix} I \\ W^\top \end{bmatrix}^\top \Phi \begin{bmatrix} I \\ W^\top \end{bmatrix} \geqslant 0.$$

Here $\Phi \in \mathbf{\Pi}_{p,\tau}$ is a given matrix.

Notation: $R(\xi) := \begin{bmatrix} -Q(\xi) & P(\xi) \end{bmatrix}, \quad v := \begin{bmatrix} u \\ y \end{bmatrix}, \quad q := m + p.$

Then $P(\sigma)y = Q(\sigma)u + w$ can be written as $R(\sigma)v = w$.

Stabilization using input-output data

all systems consistent with the data

(Unknown) coefficient matrix of $R(\xi)$: the $p \times qL$ matrix

$$\tilde{R} := \begin{bmatrix} -Q_0 & P_0 & -Q_1 & P_1 & \cdots & -Q_{L-1} & P_{L-1} \end{bmatrix}$$

Arrange the data $u(0), u(1), \ldots, u(T), y(0), y(1), \ldots, y(T)$ into the vectors

$$v(t) = \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}.$$

(Adapted) Hankel matrix asociated with the data:

$$H(v) := \left[\begin{array}{cccc} v(0) & v(1) & \cdots & v(T-L) \\ \vdots & \vdots & & \vdots \\ v(L-1) & v(L) & \cdots & v(T-1) \\ \hdashline y(L) & y(L+1) & \cdots & y(T) \end{array}\right] = \left[\begin{array}{c} H_1(v) \\ \hdashline H_2(v) \end{array}\right].$$

Stabilization using input-output data

all systems compatible with the data

Crucial observation: if a matrix $\tilde{R}$ satisfies the linear equation

$$\begin{bmatrix} \tilde{R} & I \end{bmatrix} \begin{bmatrix} H_1(v) \\ H_2(v) \end{bmatrix} = W \tag{3}$$

for some $W^\top \in \mathcal{Z}_\tau(\Phi)$ then the data $v(0), v(1), \ldots, v(T)$ could have been generated by the system $R(\sigma)v = w$.

Definition: If $\tilde{R}$ satisfies (3) for some $W^\top \in \mathcal{Z}_\tau(\Phi)$, we call the AR system with coefficient matrix $\tilde{R}$ consistent with the data.

Fact: The system $R(\sigma)v = w$ is consistent with the data if and only if

$$\begin{bmatrix} I \\ \tilde{R}^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} I & H_2(v) \\ 0 & H_1(v) \end{bmatrix} \Phi \begin{bmatrix} I & H_2(v) \\ 0 & H_1(v) \end{bmatrix}^\top}_{=:N} \begin{bmatrix} I \\ \tilde{R}^\top \end{bmatrix} \geqslant 0.$$

Stabilization using input-output data

dynamic output feedback controllers

Feedback controller for $P(\sigma)y = Q(\sigma)u + w$ of the form

$$G(\sigma)u = F(\sigma)y$$

with

$$\begin{aligned} G(\xi) &= I\xi^L + G_{L-1}\xi^{L-1} + \cdots + G_1\xi + G_0, \\ F(\xi) &= F_{L-1}\xi^{L-1} + \cdots + F_1\xi + F_0. \end{aligned}$$

Polynomial matrix of controller: $C(\xi) := \begin{bmatrix} G(\xi) & -F(\xi) \end{bmatrix}$.

Closed loop system:

$$\begin{bmatrix} C(\sigma) \\ R(\sigma) \end{bmatrix} v = \begin{bmatrix} 0 \\ I \end{bmatrix} w.$$

Goal: find a stabilizing controller, i.e., if $w = 0$ then $v \to 0$ as $t \to \infty$.

Stabilization using input-output data

informativity

Define $\tilde{C} := \begin{bmatrix} G_0 & -F_0 & G_1 & -F_1 & \cdots & G_{L-1} & -F_{L-1} \end{bmatrix} \in \mathbb{R}^{m \times qL}$.

Lemma: Let $J = \begin{bmatrix} 0 & I_{q(L-1)} \end{bmatrix}$. The autonomous system

$$\begin{bmatrix} C(\sigma) \\ R(\sigma) \end{bmatrix} v = 0$$

is asymptotically stable if and only if there exists $\Psi \in \mathbb{S}^{qL}$ such that $\Psi > 0$ and

$$\Psi - \begin{bmatrix} J \\ -\tilde{C} \\ -\tilde{R} \end{bmatrix} \Psi \begin{bmatrix} J \\ -\tilde{C} \\ -\tilde{R} \end{bmatrix}^\top > 0. \tag{4}$$

Definition: The data $u(0), u(1), \ldots, u(T), y(0), y(1), \ldots, y(T)$ are informative for quadratic stabilization if there exist $\tilde{C} \in \mathbb{R}^{m \times qL}$ and $\Psi > 0$ such that (4) holds for all $\tilde{R}$ that satisfy

$$\begin{bmatrix} I \\ \tilde{R}^\top \end{bmatrix}^\top N \begin{bmatrix} I \\ \tilde{R}^\top \end{bmatrix} \geqslant 0.$$

Stabilization using input-output data

reformulation of the problem

The Lyapunov inequality

$$\Psi - \left(\begin{bmatrix} J \\ -\tilde{C} \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \tilde{R} \end{bmatrix} \right) \Psi \left(\begin{bmatrix} J \\ -\tilde{C} \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \tilde{R} \end{bmatrix} \right)^\top > 0$$

can be rewritten as

$$\left[\begin{array}{c} I_{qL} \\ \hdashline \tilde{R}^\top \begin{pmatrix} 0 & 0 & -I_p \end{pmatrix} \end{array} \right]^\top \underbrace{\left[\begin{array}{c:c} \Psi - \begin{bmatrix} J \\ -\tilde{C} \\ 0 \end{bmatrix} \Psi \begin{bmatrix} J \\ -\tilde{C} \\ 0 \end{bmatrix}^\top & -\begin{bmatrix} J \\ -\tilde{C} \\ 0 \end{bmatrix} \Psi \\ \hdashline -\Psi \begin{bmatrix} J \\ -\tilde{C} \\ 0 \end{bmatrix}^\top & -\Psi \end{array} \right]}_{=: M(\Psi, \tilde{C})} \left[\begin{array}{c} I_{qL} \\ \hdashline \tilde{R}^\top \begin{pmatrix} 0 & 0 & -I_p \end{pmatrix} \end{array} \right] > 0.$$

Important question: When do there exist $\Psi > 0$ and $\tilde{C}$ such that

$$\mathcal{Z}_{qL}(N) \begin{bmatrix} 0 & 0 & -I_p \end{bmatrix} \subseteq \mathcal{Z}^+_{qL}(M(\Psi, \tilde{C}))?$$

Stabilization using input-output data

solution set of QMI under linear map

Important question: When do there exist $\Psi > 0$ and $\tilde{C}$ such that

$$\mathcal{Z}_{qL}(N)\begin{bmatrix}0 & 0 & -I_p\end{bmatrix} \subseteq \mathcal{Z}^+_{qL}(M(\Psi,\tilde{C}))? \tag{5}$$

Theorem[6]: Let $\Pi \in \mathbf{\Pi}_{s,r}$ and $S \in \mathbb{R}^{s\times t}$. Assume that $\Pi_{22} < 0$. Then $\mathcal{Z}_r(\Pi)S = \mathcal{Z}_r(\Pi_S)$, where

$$\Pi_S := \begin{bmatrix} S^\top & 0 \\ 0 & I \end{bmatrix} \Pi \begin{bmatrix} S & 0 \\ 0 & I \end{bmatrix}.$$

Fact: If $\Phi_{22} < 0$ and $H_1(v)$ has full row rank then $N_{22} < 0$.

Under these assumptions, (5) is equivalent to $\mathcal{Z}_{qL}(N_S) \subseteq \mathcal{Z}^+_{qL}(M(\Psi,\tilde{C}))$ with $S := \begin{bmatrix}0 & 0 & -I_p\end{bmatrix}$.

Matrix S-lemma leads to conditions for stabilization using input-output data[7].

[6] van Waarde *et al.*, "Quadratic Matrix Inequalities with Applications to Data-Based Control", *SICON*, 2023.

[7] van Waarde *et al.*, "A Behavioral Approach to Data-Driven Control With Noisy Input–Output Data", IEEE TAC, 2024.

Conclusions

Conclusions

General framework of data informativity:

1. Set of systems consistent with the data
2. Control all systems in this set

Matrix versions of Yakubovich' S-lemma

Conditions for data informativity for

- stabilization
- $\mathcal{H}_2$ control
- and many others...

Control design via linear matrix inequalities

- Controllers are obtained directly using a finite batch of data
- Non-conservative design and reduction of variables

Ongoing work: experiment design for data-driven control

Thank you!